

multiples au cours d'examens de politiques institutionnelles, peu d'exemples abondent en ce sens. Cet article décrit un processus de révision de la politique d'évaluation institutionnelle impliquant l'utilisation d'un questionnaire afin de connaître les points de vue de 269 étudiants sur l'ébauche d'une politique. Parmi les conclusions principales, on compte un manque d'orientation pour utiliser l'évaluation qui complétera la formation, ainsi qu'un manque de clarté quant aux buts de l'évaluation. En outre, la politique finale semblait manquer d'orientation quant à l'évaluation en tant que soutien à l'apprentissage et à la formation instructive, bien qu'on mise énormément sur le rôle de l'évaluation dans la mesure de la réussite malgré l'emphase que les étudiants mettent sur ces deux dernières caractéristiques. Les résultats de l'étude pointent vers d'importantes contributions théoriques que les étudiants de l'étude apportent aux examens des politiques institutionnelles, et vers les défis pratiques que les institutions doivent affronter pour fournir des mécanismes qui facilitent l'engagement et reflètent des changements culturels.

Introduction

It is not surprising that advances in promising teaching and learning practices provide the impetus for many institutions across the globe to revise their policies, in a desire to maintain alignment with the emerging higher education literature. Recent advances in the field of student assessment and evaluation have highlighted the role of assessment practices in the learning process. These revisions are encouraging and reflective of a contemporary view of assessment, yet many of the efforts and literature have focused on *what* and *how* during processes of implementation rather than on *who* and *why* for consumers of the information. That is, during a review of institutional policies, there appear to be limited efforts to consult with the consumers of the assessment policies and practices, for whom the largest population are often undergraduate students. Additional perspectives—for example, of graduate students, faculty, and administrators—would also contribute to a more comprehensive assessment policy, but this study is limited to a focus on the undergraduate perspective. To begin to address this issue, we provide an account of an institutional assessment policy review process that sought 269 undergraduate students' perspectives on a draft document. The results were intended to provide advice on revisions to policy that related to assessment purposes and principles.

Assessment practices across educational and employment contexts have been in the midst of a paradigm shift (Darling-Hammond & Bransford, 2007; Shepard, 2000). This shift from a *culture of testing* towards a *culture of learning* is occurring in response to an emerging understanding related to how the learning process occurs and how assessment can support classroom practices. In this respect, the most influential literature has been discussing the impact of assessment practices on learner motivation, and the need for broadened assessment practices beyond the purpose of measuring achievement, to include assessment as supportive of student learning as well as informative to the instructional process. This can be translated into higher education classroom practice in various ways, such as embedding assessments within the instructional process and implementing more authentic assessments. Assessments with greater authenticity are those that reflect real-life skills and in so doing provide students with access to feedback that is relevant for further

developing intended skills. It is hoped that by providing this access, students will be better able to address their own weaknesses (Joughin, 2009). Thus, students become one of the primary consumers of the assessment information yet sometimes lack the necessary experience to recognize the benefits of feedback (Price, Handley, Millar, & O'Donovan, 2010).

The need for greater embedded opportunities for assessment within the instructional process is well established in the current higher education literature related to assessment and course design. Further, there is a need to consider how assessments are aligned with both learning outcomes and instruction. It is therefore important to clearly identify, as instructors, what we want the students to learn, and to design our instruction around helping them learn it—including by using formative assessment (i.e., assessment in which grades are assigned for the purpose of informing learning and instruction)—then assess how well they learned it using summative assessment (i.e., assessment in which grades are assigned for the purpose of measuring learning). This perspective aligns with what Suskie (2009) calls a contemporary approach to assessment, which contrasts with a traditional approach; these approaches are summarized in Table 1. As authors, although we appreciate the dichotomy that Suskie presents, we also acknowledge that this might over-simplify the construct and undermine the issues involved in the complex construct of assessment. Still, we conceptualize it as a useful starting point from which to discuss classroom assessment practices.

Table 1.
Comparison of Contemporary and Traditional Approaches to Assessment

Contemporary	Traditional
Aligned with learning goals	Planned and implemented without consideration of learning goals
Focused on high-order thinking and performance skills	Often focused on lower-order thinking skills
Developed based on current research related to teaching and assessment	Often of poor quality because instructors have lacked opportunities to learn about high-quality assessment practices
Used to improve teaching and learning as well as to evaluate and assign grades	Used only to evaluate and assign grades

Note. Modified from Table 1.2 of Suskie (2009).

While more contemporary assessment practice should now involve collecting information on student achievement and performance through the use of a variety of tasks designed to monitor and improve student learning (Gipps, 1994), actually changing classroom practice is difficult because assessment must now perform several tasks at once (Boud, 2000; Ramsden, 2003). Among the challenges for institutional policy-makers is revising the existing assessment policy to reflect contemporary assessment practice and then providing guidance to instructors regarding how to implement these changes with fidelity within the instructional environment.

One of the greatest hindrances to achieving change in assessment practices within the higher education context is resistance among instructors (Deneen & Boud, 2014). Within

assessment practices, in addition to the call for alternative methods, student engagement has become a rapidly growing notion in higher education organization and management (Leach, 2012).

Limited research exists on effective ways of engaging students in higher education institutional processes. In many institutions, policies are in place to guide student participation in the governance structures, yet mechanisms for their voices to be heard may not be effective. One of the challenges may be the institutional culture in which the policies are enacted; for example, “[m]ost of student engagement research focuses on how students interact with their educational environment, rather than the way the institutional environment engages with them” (van der Velden, 2012, p. 229). The potential for enhancing students’ learning experience through student participation in assessment policy decisions is worthy of consideration because of the well-established connection between assessment experiences and student motivation to learn, yet there is a dearth of examples for how to undertake this task.

Scotland’s pioneering approach to student engagement provides a systems-level example with great potential for guiding practices within policy making to impact student experience. Specifically, the aim of the approach is to put “students at the heart of decisions about quality and governance” (Student Participation in Quality Scotland [sparqs], 2013a, p. 4). Among the five key elements of student engagement pertinent to the current student are roles for students and institutions. For the former, this specifically refers to students engaging in their own learning and students working with their institution to shape the direction of learning; for the latter, this refers to the institution providing formal mechanisms for quality and governance. Not surprisingly, a culture of engagement and of valuing the student contribution has emerged as among the six features of effective student engagement.

Providing a mechanism for students to review the draft policy document was intended to be a first step towards enhanced mechanisms for engaging students in policy decisions. Indeed, undertaking the policy review process was seen as an opportunity to engage in an “inclusive conversation about assessment and grading, and to come to a consensus on both the purposes of assessment, and principles surrounding assessment, that would govern university-level policy” (Luth, p. 2). It was a focus on the student learning experience, including how feedback was being provided and grade assignment was *actually* occurring in practice, that provided the impetus for the review of institutional assessment policy which was undertaken by a subcommittee of the institution’s Committee on the Learning Environment. The following contextual information is provided so that readers may apply some of the activities to their own contexts.

Study Background

The study took place at a large, research-intensive university in Western Canada. Each author of this paper made a unique contribution to the study, which can largely be attributed to their differing roles during the review process; one author (Luth) was the chair of the Subcommittee on Assessment and Grading (2010–2011), another (Poth) was a member of this subcommittee, and the third (Riedel) was a research assistant to the other authors. As chair, Luth had the initial tasks of recruiting members who represented the diverse roles of those involved in student assessment and then delineating the subcom-

mittee's terms of reference. As a member, Poth contributed expertise in assessment and measurement and led the research initiative. As a research assistant, Riedel brought expertise in quantitative data analysis and organizing data collection.

The subcommittee met monthly in 2010, during which the chair facilitated several activities that culminated in a report; this he intended as "the beginning of a conversation in the academy, not the end of one" (Luth, 2010, p. 5). Indeed, the subcommittee was guided by the idea that "change depends on generating consensus on principles rather than prescribing specific practices" (Joughin, 2009, p. 5). One of the strengths of the committee membership was its inclusion of multiple perspectives on assessment, including students' and non-academic staff's (i.e., from offices of the Registrar and Student Ombudservice) in addition to academic staff's. A further strength was the focused terms of reference agreed upon by the subcommittee members (Luth, 2010):

- Survey expectations and experiences of students, instructors and administrators.
- Review recent literature on effective/best/exemplary practices.
- Identify examples/stories where assessment supports excellence in learning and teaching at the University of Alberta and how they came about.
- Formulate recommendations.

To that end, the chair facilitated discussions in which consensus was reached related to two questions: Why do we assess? Are there principles of fair and appropriate assessment on which we can agree? These discussions were informed by information related to what was currently (a) happening across similar institutions at the organizational level, (b) emerging in the literature that should be seen as guiding practice within post-secondary contexts, and (c) being enacted as practices on our own campus. The answers to these questions from the subcommittee were anticipated expected to inform the framing of a university-wide policy and were presented as part of a report that included a draft document of six assessment purposes and six principles (see Table 2). This report was then circulated across campus as well as made public, and the subcommittee made explicit its desire to continue consultations with students, faculty, and administrators. In the following academic year, a new committee was tasked with continuing the work, and subsequently, members of the Academic Standards Committee were responsible for formulating the final policy document.

Study Purpose

The overall goal of the study was to incorporate student perspectives during the review of institutional policy, which was anticipated to inform revisions to assessment policy, and then to provide an account of this process. To do this, we first present the empirical findings from a sample of undergraduate students' perspectives on the draft purposes and principles as well as the challenges that they experienced related to assessment at the university. We then provide a description of the extent to which the learner perspective is reflected in the final university-wide assessment policy.

Table 2.

Draft Purposes and Principles from the 2010 Subcommittee on Assessment and Grading

Purposes of Assessment	Principles of Assessment
To evaluate – should produce a judgement about the student's achievement of the learning goals/outcomes of the course.	Should be integrated into and aligned with the learning experiences and intended outcomes of a course.
To rank students – for scholarships and advancement (e.g., entry into graduate or professional programs).	Must validly and reliably measure expected learning outcomes, both disciplinary content and higher-order outcomes.
To communicate – the grade in the end is all the outside world will know (or, perhaps, all the student will remember) about their achievement in that course.	Should build students' ability to self-assess and self-reflect, and promote deep learning.
To improve – both learning on the part of the student and teaching on the part of the instructor.	Should involve varied assessment strategies, as appropriate for the subject.
To motivate – there is general agreement in the literature that assessment drives student learning: what they study, what they focus on, how they approach their learning.	Should include early opportunities for students to align their understanding of expectations on assignments with those of the instructor.
To encourage self-assessment and reflection on learning by the student.	Must be transparent.

Method

The current study took place during an institutional assessment policy review process (January 2010 to July 2012). The study itself was conducted during September–November 2010, after having received ethical clearance. First, a survey design was used to explore the perspectives of 269 undergraduate students who completed an online questionnaire. Then, the results of the questionnaire analysis were disseminated to the subcommittee members and, subsequent to its public release, the final policy document was reviewed.

Student Perspectives Explored Using a Survey Design

A survey design was chosen because of its appropriateness for generating information related to patterns and trends associated with a selected population (Creswell, 2009). In the current study, the focus was on exploring the perspectives of undergraduate students on the draft assessment purposes and guiding principles, as well as documenting the assessment challenges they had experienced within the higher education context. Details related to participants and recruitment as well as to the development, administration, and analysis of the online questionnaire are presented below.

Participants and recruitment. Participants represented a convenience sample of 280 undergraduate students, recruited via an undergraduate participant pool from two second-year courses who received course credit for their participation. Participants indicated their interest in the present study by sending a research assistant an email and were subsequently sent the link to the web-based questionnaire in October 2010. A reminder

was sent two weeks later, and of the 280 who were sent the link, 269 completed all sections of the questionnaire (96% response rate). The majority of participants were enrolled in programs within the Faculty of Education (86%), and 78% self-identified as female. The participants' mean age was 23 years, and the median age was 21 years. Thirty percent of the students indicated that they intended to pursue graduate studies. Our population statistics tell us that this sample appears to be representative of the gender ratio and age of our population of students in the program (76% female; median age 21.6).

Online questionnaire. A questionnaire was used because of its usefulness as a cost-effective and efficient data source for accessing the perspectives of a selected population (de Vaus, 2001). The questionnaire generated both quantitative and qualitative data. The first two sections involved researcher-created items based on the original statements from the policy's draft purposes and principles. The first section involved 12 items to access students' perceptions of the appropriateness of purposes, using a five-point Likert rating scale anchored at the endpoints (1 = "not appropriate at all," 5 = "very appropriate"). The second section involved 22 items to access students' perceptions of their agreement with the guiding principles, again using a five-point Likert scale anchored at the endpoints (1 = "strongly disagree," 5 = "strongly agree"). The third section involved a researcher-created open-ended question related to students' assessment experiences: What are the major challenges related to assessment that you have experienced? Respondents were provided unlimited space in which to answer this question. The final section included four demographic items related to faculty, gender, year of birth, and intention to pursue graduate studies.

The development of the questionnaire involved three phases. First, the research team sought the assistance of a measurement expert (T. Rogers) to translate the original statements from the draft purposes and principles into questionnaire items that would meet the guidelines for high-quality items. Among the changes were to break the statements into smaller parts so that they were no longer double-barrelled (Nunnally & Bernstein, 1994). For example, an original statement, "The purpose of assessment is to improve—both learning on the part of the student, and teaching on the part of the instructor," became two items: "Assessments are appropriate for improving the quality of teaching I receive" and "Assessments are appropriate for improving my learning." A similar process was undertaken for the items included in the guiding principles section. Second, the research team used think-aloud protocols (Willis, 2005) with two undergraduate students to inform clarity of instructions and items. Finally, a panel of experts with experience in higher education assessment reviewed the questionnaire and rated the fit between items and the purposes and guiding principles to which the items were referenced. This feedback led to modification of a few items and the wording of instructions.

Following administration of the survey, to reduce the complexity of the item sets for the first two sections of the questionnaire, factor analysis was applied separately to the purposes of assessment and the guiding principles of assessment. First, a principal components extraction was performed to identify the number of factors. Application of the Kaiser-Guttman rule of eigenvalues greater than one, and Cattell's scree test, suggested four factors for the purposes of assessment and five factors for the guiding principles of assessment. Second, principal axis factoring with oblique (direct oblimin) transformation was employed to obtain a factor pattern that exhibited a simple structure and that was interpretable for each section.

For the third section, related to challenging assessment experiences, Poth and Riedel independently undertook the inductive analysis of students' responses, using a constant comparison method (Charmaz, 2006). This approach to coding employs constant comparison and memoing and results in themes that emerge from the data. First, each researcher read the first 40 responses to the open-ended question and independently generated a preliminary list of codes while keeping track of their thoughts using memos. The researchers then compared their code lists and found a high degree of similarity across their codes; in cases of discrepancies, they discussed until a consensus was reached and a final code list was generated. Next, this finalized code list was applied by one of the researchers to the remaining responses. Then the codes were categorized to generate five themes; for example, three codes (i.e., inadequate feedback on graded assessments, lack of feedback on how to improve, few opportunities to receive feedback on graded work) were categorized into the theme "lack of feedback." Descriptive statistics were generated for each of the demographic items. A summary of each of the questionnaire sections was generated and disseminated by the subcommittee chair to its membership. The timing of dissemination was key, as it occurred while the working subcommittee group was concluding its work during the spring of 2011.

Review of Final Assessment Policy Document

The review of the final policy took place in July 2012, following its approval by the university's General Faculties Council in June 2012, to assess the extent to which the final policy document reflected the students' perspective. First, we undertook a side-by-side comparison between the principles from the final policy and the students' perspectives on the draft document and the assessment challenges they had experienced. Then we sought evidence of the extent to which the students' perspectives on the appropriateness of the assessment purposes were reflected in the final policy document.

Findings and Discussion

The findings and discussion are organized in two sections: first, the students' perspective as captured through the questionnaire findings, and then the comparison results generated by the review of the final policy document.

Students' Perspective

The students' perspective is presented in three sections related to the questionnaire: appropriateness of assessment purposes, agreement with guiding principles, and assessment challenges experienced by students.

Appropriateness of assessment purposes. A four-factor pattern emerged, accounting for 55.0% of the total variance in the section focused on the extent to which assessment purposes are appropriate (see Table 3). Factor 1 is related to enriching instructional practices (factor mean score = 3.8). Factor 2 has to do with communicating achievement information (3.0). Factor 3 is concerned with encouraging student metacognition (3.7). Factor 4 is related to supporting students' learning (4.1). It is noteworthy that the three highest factor mean scores emphasize the contemporary conception of assessment purposes as appropriate for supporting the actions undertaking by instructors (factor 1) and students (factors 3 and 4), whereas the lowest factor mean score is focused on

the traditional assessment purpose as appropriate for communicating achievement information (factor 2). Thus, our findings suggest that most students consider the summative function of assessment to be less appropriate when compared to the formative function of assessment. These findings indicate that students embrace the more contemporary conception of assessment as supporting and enhancing the teaching and learning environment in addition to simply measuring and communicating achievement information.

Also evident in our findings is the focus on instructors and students as important consumers of assessment information; students view it as appropriate (i) for instructors to use the results to inform their teaching practices, and (ii) to apply the results as active participants in their own learning process. It is important to note that the students' responses indicate a desire for greater participation, as this reflects what has been called for in the literature (e.g., Boud & Falchikov, 2006; Donald, 1997)—in other words, it is no longer sufficient for students to be passive participants. As passive participants, students are likely to view the assessment process as an activity that is done to them at the end of instruction (Boud, 2000). Instead, researchers have called for students and instructors to support learning and enrich instruction by using the information generated through embedding both formative and summative assessments within the instructional process (Suskie, 2009).

Table 3.

Item Means, Factor Means, and Factor Pattern Coefficients Extracted from Principal Axis Factoring, Applying Oblique Transformation for Items Representing the Purposes of Assessment

Items Reflecting Purposes of Assessment	Item Mean	Factors			
		1	2	3	4
Assessments are appropriate for . . .					
1. modifying my instructor's teaching.	3.8	.89			
2. improving the quality of teaching I receive.	3.9	.79			
3. enhancing the quality of my instructors' future-assessments.	3.7	.65			
4. guiding my instructor's implementation of course objectives.	3.7	.46			
5. ranking students according to their achievement of course objectives.	3.1		.92		
6. informing others about my performance.	2.8		.42		
7. encouraging reflection on my learning.	3.8			-.90	
8. encouraging my self-assessment.	3.6			-73.0	
9. motivating me.	3.8				-.72
10. improving my learning.	3.9				-.57
11. measuring my achievement of course objectives.	4.3				-.50
12. informing me about my progress.	4.2				-.36
% of total variance explained	33.3	10.7	6.5	4.5	
Factor mean score	3.8	3.0	3.7	4.1	

Note. Factor pattern coefficients with values lower than | 0.30 | are not indicated.

Agreement with guiding principles. A five-factor pattern accounting for 58.6% of the total variance emerged from the analysis of the items assessing agreement with guiding principles (see Table 4). Factor 1 has to do with the alignment of assessment with instruction and course outcomes (factor mean value = 4.2). Factor 2 is related to the transparency of the assessment process (4.3). Factor 3 is concerned with the consistency of grading practices (4.6). Factor 4 is related to the ability to assess higher-order cognitive skills (3.9). Factor 5, consisting of two items, is concerned with awareness of the assessment criteria (4.8). It is notable that the two highest factor mean scores emphasized the need for assessment criteria to be accessible to students (factor 5) and grading practices to be uniform across differing contexts (factor 3), whereas the lowest factor mean score focused on assessing higher-order cognitive skills (factor 4). These findings might indicate that students consider it their right to have grading procedures that are both valid (i.e., reflective of the communicated assessment criteria) and reliable (i.e., consistent across markers and terms), and that they demand greater access to this information. These ideas are consistent with what are considered fair assessment practices in the literature (Dochy 2009; Shepard 2000) and possibly reflect a shift towards students becoming more active and informed participants in the assessment process.

Table 4.
Item Means, Factor Means, and Factor Pattern Coefficients Extracted from Principal Axis Factoring, Applying Oblique Rotation for Items Representing the Guiding Principles of Assessment

Items Reflecting Guiding Principles of Assessment	Item Mean	Factors				
		1	2	3	4	5
Assessment should be integrated with the course outcomes.	4.2	.80				
Assessment should be aligned with the course outcomes.	4.2	.75				
Assessment should be aligned with instruction.	4.2	.72				
Assessment should be integrated with instruction.	4.2	.67				
Assessment should be linked to the intended course outcomes.	4.1	.62				
Assessment should be guided by a clearly articulated policy at the department/faculty level that is consistent with university policy.	4.0	.43				
Instructors should discuss with their students the appeal procedures for course grades.	4.2		.91			
Students should be made aware of the appeal procedures for individual assignments.	4.4		.89			
Students should be made aware of the appeal procedures for course grades.	4.4		.89			
Instructors should discuss with their students the appeal procedures for individual assignments.	4.2		.88			

Items Reflecting Guiding Principles of Assessment	Item Mean	Factors				
		1	2	3	4	5
Instructors should include opportunities for students to align their understanding of assessment criteria with that of the instructor.	4.3		.43			
Instructors should use varied assessment strategies, as appropriate for the intended course outcomes.	4.5			.32		
Instructors should be consistent in their grading within a course.	4.7				-.91	
Instructors should be consistent in their grading across multiple sections of the same course.	4.4				-.82	
Instructors should be consistent in their grading across different terms of the same course.	4.6				-.60	
Assessment must be representative of the intended course outcomes related to higher-order thinking.	3.8					.86
Assessment must consistently measure higher-order thinking.	3.6					.85
Assessment should be related to higher-order thinking.	3.9					.70
Assessment should enhance students' ability to develop discipline-specific expertise.	3.9					.39
Assessment should develop students' ability to self-assess.	4.0					.37
Students should be made aware of the assessment criteria for individual assignments.	4.8					-.83
Students should be made aware the assessment criteria at the beginning of the course.	4.8					-.75
% of total variance explained		26.7	15.9	7.1	4.6	4.3
Factor mean score	4.2	4.3	4.6	3.9	4.8	4.2

Note. Pattern coefficients with values lower than | 0.30 | are not indicated.

Assessment challenges experienced by students. Five themes emerged from the analysis of the reported assessment challenges: unclear expectations, limited strategies, missing feedback, unfair grades, and poor quality (see Table 5). Unclear expectations emerged with the greatest frequency, with an emphasis on specific individual assessments and the course in general. Students expressed frustration when information related to the expected outcomes was not communicated and when the assessments did not reflect the content that had been emphasized during instruction: “When the instructor tells you that the readings from the text are of minimal importance and to focus on the notes, and then the majority of the assessment is from the text.” The impact on students’ learning experience of such instructional practices is highlighted by one student who wrote: “Unclear or vague

instructions on assignments cause the students to miss the point of what the instructor was looking for." Similar frustration referred to students' confusion about what to focus on within a course—for example, when the professor did not provide an overall orientation as to the course purpose and the instructor's expectations. To enhance instructors' communication of their expectations, several students suggested providing access to sample questions or assignments. These findings clearly point to the need for instructors to actively engage students in ongoing opportunities to clarify their understandings of assessment expectations. Indeed, these practices are supported by literature pointing to effective strategies for enhancing the communication of expectations, including discussing the expectations and providing exemplars of questions and/or assignments (Dysthe, Engelson, Madsen, & Wittek, 2008).

Table 5.

Themes, Sub-theme Frequencies, and Representative Quotes from Inductive Analysis of Students' Assessment Challenges

Theme	Sub-theme	Freq.	Representative Quote
Unclear expectations	Specific to individual assessments	59	"When instructors don't communicate details on how they want their exams/assignments and expect you to be able to fill in and know what they want."
	In general for the course	40	"Teachers not effectively communicating the desired outcome."
Limited strategies	Limited type	40	"Many courses only give one type of assessment."
	Frequency of administration	7	"I find so often you have the midterm and final and that's all the assessment you get."
	Level of cognition	4	"Most tests emphasize memorization with little focus on understanding."
Missing feedback	In general during the assessment process	21	"I experience difficulty in assessment when I don't receive feedback regarding assignments or exams."
	Specific for an assigned grade	13	"Confusion as to why a certain mark was received because work wasn't marked or no comments."
	For the purpose of improvement	13	"Not given any feedback on assignments. It is difficult to know what I need to work on."
Unfair grades	Inconsistencies across markers	17	"It is often frustrating when multiple graders are not standard across a class."
	Norm-referenced curving	8	"Being put on the curve so that even with a high percentage grade, you end with a low letter grade."
Poor quality	Unclear wording	10	"Sometimes questions are written unclearly, and if I get it wrong it's a matter of wording, not that I don't know the material."
	Lack of discriminating options	6	"Telling the difference between the right and the best right answer."

Students reported that limited strategies are used within their post-secondary classroom environments in terms of type of method, frequency of administration, and level of cognition. The over-reliance on particular types of assessment strategies—that is, exams in general—was of particular concern to these students. Exams contrast with presentations or written assignments in that they are typically timed and invigilated. Specifically, students were concerned with the prevalent format of exams being multiple choice (also known as selected response). This concern was attributed to individual students' abilities, preferences, and experiences; while some students might favour multiple-choice exams, others might be disadvantaged by not being able to express themselves differently. One student noted:

The types of assessments offered, because some students are better at certain assessments than others. So if only one type of assessment is offered for the whole course (ex: exams), it might affect the marks of some students, because they may have difficulty with exams, but are very good at presentations or assignments.

In addition to a lack of diversity in current implementations of assessment strategies, students also pointed to the need for greater frequency of assessments as well as an increased focus on higher-order cognitive skills. Students specifically associated an increased number of assessments administered with a reduction in the pressure they felt when only a few assessments contributed to their course grade. Indeed, they considered that more frequent assessment opportunities would yield a more accurate representation of what they had learned in the course, because of the distorted impact that a low grade has on a course grade when a few assessments are heavily weighted: "No chance to make up a bad grade, especially in courses with few assessments." A number of students further emphasized that multiple choice exams focus on assessing students' ability to recall specific details (i.e., lower-order cognitive skills) rather than their grasp of big concepts. This was of particular concern when an exam covered a wide range of material: "When exams were more focused on memorization of textbook points rather than concepts and theories." Together, these findings point to the need for instructional practices that integrate a greater diversity of assessment strategies, more frequent assessment opportunities, and strategies that also target higher-order thinking skills. The positive impacts of such assessment practices on learning are well documented in the literature (Suskie, 2009).

Missing feedback emerged as a key assessment challenge for students, related to three aspects: (a) in general during the assessment process, (b) specific to an assigned grade, and (c) for the purpose of improvement. The majority of students referred generally to the lack of available information during the assessment process. Of particular note were the students who reported receiving inadequate feedback to justify an assigned grade: "When assignments are returned, having no feedback or reason for the grade I received." Students also expressed a desire for timely and specific feedback that provided guidance for improvement: "When I am not given any feedback on assignments it is difficult to know what I need to work on." One student highlighted the inadvertent repeated errors he/she made when assessed work was not returned in a timely manner or not returned at all: "[O]ften I have handed in several [assignments] before I get the first one back. This is a huge problem as I could have made a simple mistake in the first, that I repeated in the following assignments." Together, these findings clearly indicate the powerful influence

of timely and specific feedback to optimize its usefulness for supporting learning (Hattie & Timperley, 2007). In addition, our findings align with research demonstrating that providing grades accompanied with written comments is consistently more effective than simply providing grades on assessed work (Black & William, 1998; Crooks, 1988).

Students raised two issues related to unfair grading practices they had experienced: inconsistencies among markers (i.e., teaching assistants and instructors in the same course) and norm-referenced curving (i.e., when student grades are changed, often to maintain a predetermined course average). Many students expressed their frustration over inconsistencies among multiple markers: "It is often frustrating when multiple graders are not standard across a class." Notably controversial from the student perspective are curving practices: "In classes where the majority (i.e., 20 of 25) are doing very well, and if it's graded on the 'curve', someone has to get a 'C', which is unfair." Students, because of the resulting distortion to their course grades, viewed both of these practices as unfair. Similar to our findings, several researchers (Ecclestone & Swann, 1999; Morgan, Dunn, Parry, & O'Reilly, 2004; Weigle, 1999) point to the need for fair grading practices; among those highlighted are implementing marker reliability training, communicating historical grade ranges, and assessing according to articulated intended student outcomes (i.e., criterion-referenced grading rather than norm-referenced grading).

Finally, students reported challenges associated with poor-quality assessments, specifically identifying unclear wording and the lack of discriminating options. Concerns related to the former were focused on written questions that were confusing and hence ultimately distorted course grades: "Knowing the material but the wording of questions makes it difficult to apply my knowledge no matter how hard you study. . . . [But] you still know the material." In a related concern, students highlighted the challenges associated with differentiating among multiple choice options that offer little or no basis for discrimination: "Telling the difference between the right and the best right answer." These findings underscore the pressing need for greater expertise in developing high-quality assessments, which aligns with the results of several studies pointing to common problems within higher education assessments (e.g., Albanese, 1993; DiBattista & Kurzawa, 2011). Similarly, Terrant and colleagues (2006) found that almost half of the multiple choice questions used in a nursing program violated guidelines for writing multiple choice items. If students are to be accurately and fairly assessed, then the assessments must be of high quality (Brookhart, 2011).

Review of the Final Policy

The review of the final policy document is presented in two sections: comparison with draft principles and challenges experienced, and comparison of draft purposes with students' perspectives.

Comparison with draft principles and challenges experienced. Evidence of similarities and differences emerged from comparing the final policy principles with the students' perspectives on the draft principles and the assessment challenges they experienced (see Table 6 for a summary). Specifically, similarities exist between the policies relating to the first principle—promoting the alignment of assessments—and among the policies and challenges with respect to the principles emphasizing communication and transparency of assessment methods, standards, criteria, and processes (i.e., the third,

fourth, and seventh principles), and the need for embedded and timely assessment-related processes (i.e., the sixth principle). Differences exist in terms of what the final policy principles suggest with respect to developing innovative assessments (i.e., the second principle), the measurement properties of assessments (i.e., the fifth principle), and a lack of reference to feedback in the final policy.

Table 6.

Summary of the Comparison of Final Policy Principles with the Students' Perspectives on the Draft Guiding Principles and Assessment Challenges Experienced

Final Policy Principles	Students' Perspectives on...	
	Agreement with Draft Principles	Assessment Challenges Experienced
1. Assessment should be integrated into and aligned with the learning experiences and stated objectives/outcomes of a course and program	Alignment of assessment with instruction and course outcomes	
2. While this policy sets out the minimum expectations concerning the design and delivery of assessments, it does not limit the development of other, additional, innovative forms of effective assessment, provided they are compatible with the principles stated in this policy.		Limited strategies (limited type)
3. General assessment methods and grading standards must be communicated clearly to students at the beginning of the course or program of study.	Transparency of the assessment process	Unclear expectations (course)
4. Clear and transparent assessment criteria should be provided to students throughout the course.	Awareness of assessment criteria	Unclear expectations (assignments)
5. In assessment, the University is committed to providing reliable and valid information in which students, prospective employers, and accrediting bodies can have confidence.		Poor quality
6. Where possible, assessment should be multifaceted (varied) and timely. Student achievement and performance should be assessed in a formative manner during a course and in a summative manner both during and at the end of a course and program.	Assess high-order cognitive skills	Limited strategies (frequency of administration and level of cognition)
7. In the design, delivery, and reporting of summative assessments, the University is committed to open, accountable, and equitable processes.	Consistency of grading practices	Unfair grades

The first principle from the final policy includes similar wording to that of the draft principle—for example, integrated into and aligned with instruction (i.e., learning experiences) and course outcomes (i.e., stated objectives). It is interesting to note that students did not report any challenges with respect to lack of alignment. This may be attributed to students' lack of understanding of *how* assessments should be integrated within the instructional process.

The third, fourth, and seventh principles represent a shared emphasis on maintaining transparency in assessment methods, standards, criteria, and processes by communicating assessment information, yet the final policy document differed in the extent to which it provided useful implementation guidance. Specifically, the scope of the third principle was limited to the instructor communicating about assessment methods and grading standards at the beginning of the course. As a result, the final policy lacked information on how grading consistency among multiple markers would be maintained, although students had highlighted this as a challenge they had experienced. The fourth and seventh principles represent the common focus on transparency throughout the assessment process. Specifically, both the fourth principle in the final policy document as well as the draft principles highlighted the need for students to be *provided* with ongoing access for clarifying assessment criteria. It appears that the instructor is thereby responsible for providing these opportunities for both the course and the assignments, which begins to address challenges experienced by students. This contrasts with the seventh principle in the final policy document, where the university appears to be responsible for upholding the integrity of summative assessments. In many ways these findings do not assign the student the greater participatory role that has emerged, whereby students desire to be the *who* of assessment.

The sixth principle, which focused on assessment-related processes that are both varied in method and timely in use, not only addressed one of the major challenges reported by students (i.e., use of limited types of strategies) but also specified the use of both formative- and summative-type assessments that had not been previously attended to in the draft document. It is noteworthy that the final policy no longer reflected any reference to level of cognition; specifically, greater use of higher-order thinking skills was highlighted in the draft document and as a challenge experienced by students. The principle's focus on embedding assessment processes is encouraging—yet in many ways, to be useful for guiding practice, the *what* (i.e., the breadth and depth of information) to be shared must be clearly articulated.

The second and fifth principles in the final policy document make unique contributions to the *how* of implementation, which is not addressed by either the draft or the final policy document; the former speaks to developing innovative assessments, the latter to enhancing measurement rigor. Promoting the idea that ongoing innovations in assessment are necessary is essential for maintaining relevance with emerging learning environments. For example, no topic has become more central to innovation and practice in educational assessment than computers and the Internet (e.g., Ricketts & Wilks, 2002; Thelwall, 2000). While the fifth principle specifies "reliable and valid information," which can be interpreted to encompass the challenges identified by students related to unclear wording and lack of discriminating options, its purpose is to apply measurement properties. The difficulties for implementation are that the terms are not defined in any way and would require a degree of understanding of measurement and *why* it was necessary to meet these standards.

Finally, although the principles in the final policy have addressed most of the assessment challenges reported by students, one exception remains, related to the role of feedback. It should be noted that feedback is represented within the introductory statement: “It [assessment] is undertaken in a formative manner to provide feedback to students.” Yet it remains unclear whether feedback is also expected as part of graded (that is summative) assessments. This gap across the policies may represent an enduring lack of focus on the student as the *who* of assessment, and on the *why* of assessment—that is, from the students’ perspective of assessment’s broadened purpose being to support learning as well as to measure achievement and inform instruction.

Comparison of draft purposes with students’ perspectives. Whereas our study sought the extent to which each of the purposes from the draft document was perceived by students to be appropriate, the final policy simply alluded to the multiple purposes. Indeed, both formative and summative purposes are reflected in these statements: “Assessment . . . is undertaken in a formative manner to provide feedback to students and in summative form to measure the level of student achievement . . . achievement is communicated to a variety of stakeholders.” Therefore, two of the purposes from the draft are clearly represented in the final document—that is, to communicate achievement information and support student learning. Also clear is that the final two purposes, encouraging student metacognition (i.e., self-assessment) and enriching instructional practices, are missing from the final policy. These are important considerations, given that Boud and Falchikov (2006) argue that the development of skills such as self-assessment is crucial in today’s rapidly changing society, which requires its members to be lifelong learners. Moreover, the policy is silent on one of the essential purposes of the current assessment understanding, namely, the use of assessment results for instructional quality improvement. The intended audience of assessment information—the *who* of assessment—appears to be somewhat the student but not at all the instructor.

Implications, Limitations, and Future Directions

This study has two important implications for informing how a review of institutional policy can be undertaken to generate more learner-centred policies and practices. First, the study provided an illustrative example of a process that was cost-effective and time-efficient in seeking the perspectives of undergraduate students. However, the study may be limited by the methods used for sampling and for data collection. The former is related to the use of convenience sampling, which limits the generalizability of the study findings to larger undergraduate populations. Given the over-representation in the study’s sample by a single faculty and the female gender, it would be imperative for a follow-up study to confirm findings across faculties on campus. This is because students enrolled in the teacher education program are more likely to have taken an assessment course and as a result may be more attuned to high-quality assessment practices. The latter (the data collection method) is related to the study’s use of a questionnaire with only students, which limited our ability to probe written responses and made the students’ perspectives the sole focus rather than also capturing the instructors’ perspectives as well as the experiences of the subcommittee members. Further research is needed to: (a) address the methodological issues by using additional data sources and (b) replicate this study for greater generalization and understandings of the learner’s perspective on assessment policies and practices.

Second, by beginning to addressing the need for students to play a more active role in their assessment process, the results captured students' perspectives on the draft document outlining purposes and guidelines, as well as the challenges students experienced. We have thereby learned that these students want the assessment process to reflect: (a) multiple purposes within a contemporary view of assessment, including supporting learning and informing instruction, as well as measuring achievement; (b) the principles of fair assessment, including the use of diverse strategies that are valid and reliable; (c) the role of students as primary consumers of assessment information who, as such need, to be informed about expectations and grading practices; and (d) assessment's function in enhancing the teaching and learning environment, including the role of instructors to use the information to enhance their instruction and provide feedback that is timely, relevant, and specific for guiding students' learning. Understanding the role that students played in the creation of the final policy document is limited by not being privy to either the subcommittee's review of the findings from the questionnaire or the processes they undertook to finalize the document. This lack of access was unanticipated yet attributable to both committee membership changes (which is usual practice from year to year, to enable a greater number of participants) and a lack of ethical clearance to talk to these committee members. We therefore were unable to assess the extent to which the student contribution was valued; this is one of the effective features of student engagement highlighted in the student engagement framework for Scotland (sparqs, 2013b).

In the future, it would be prudent to further explore committee members' reactions to student input and ascertain what actions they undertake to integrate these student perspectives into their policies and practices. This inaccessibility is one example of the many challenges experienced within the higher education context, which limits the transparency of the process, making it difficult to assess the reasons for revisions from draft to final document. Although frustrating, this lack of insight into the underlying process is not unique; indeed, Daneen and Boud (2014) during their attempts to enact changes in assessment practices noted: "the results from a change attempt are often quite different than what was intended" (p. 577). Among the reasons noted when various stakeholders are involved are differences in interpretations, and the extent to which stakeholders ignore or dismiss changes (Trowler & Bamber, 2005).

Conclusion

This study highlights the contributions that students can offer to institutional policy reviews and the challenges institutions face when seeking input from multiple perspectives. It is encouraging that students in the current study expressed their desire as learners to be considered primary consumers of assessment information that is relevant to supporting their learning as well as measuring it. The institutional review process was successful in that the final policy document does reflect many of the purposes that the students considered appropriate, yet what remains to be included is how instructors can also use assessment information to enhance the teaching environment (e.g., through providing students with feedback).

The processes involved in changing policies within the higher education context present a challenge for institutions, and the present study demonstrates that assessment policy reviews are by far no exception. The call for more active participation of students in policy

changes has met with mixed results, largely attributable to lack of institutional cultural shifts, which would promote further mechanisms for students to actively contribute to changes. Being uncomfortable with making changes to their roles so as to provide mechanisms for shifts in culture. As students are increasingly positioned as consumers, institutions need to improve the extent to which these consumers' demands are met (Furedi, 2010). Indeed, some academics might feel that their function is being reduced to that of a service provider (Deneen & Boud, 2014) and that students may not have the assessment expertise to suggest innovative practices. This study aligns with Bevitt's (2015) solution: "The argument for more student involvement may assume greater legitimacy if a student experience approach to assessment is understood not as a replacement to other assessment roles but as an accretion" (p. 115). The current study contributes important insights for subcommittee members to consider when undertaking a review of institutional policies, including (i) the value of integrating multiple perspectives and (ii) the need for assessing the alignment between aspects of a policy document, to include both the *what* and the *how* of implementation as well as the *who* and the *why* of consumers. 

Acknowledgements

This work was supported by funding from the Office of the Provost and Vice-President (Academic), and by a Support for the Advancement of Scholarship grant from the University of Alberta, received for "Investigating the Influence of Assessment Practices on the Undergraduate Learning Environment: The Student Perspective." The authors would also like to thank Todd Rogers and Erin Sulla for helpful comments during the review of a draft of this manuscript.

References

Albanese, M. A. (1993). Type K and other complex multiple-choice items: An analysis of research and item properties. *Educational Measurement: Issues and Practice*, 12(1), 28–33. doi:10.1111/j.1745-3992.1993.tb00521.x

Bevitt, S. (2015). Assessment innovation and student experience: A new assessment challenge and call for a multi-perspective approach to research. *Assessment & Evaluation in Higher Education*, 40(1), 103–119. doi:10.1080/02602938.2014.890170

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167.

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399–413. doi:10.1080/02602930600679050

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.

Charmaz, K. C. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Creswell, J. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481.

Darling-Hammond, L., & Bransford, J. (2007). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.

Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment and Evaluation in Higher Education*, 39(5), 577–591. doi:10.1080/02602938.2013.859654

de Vaus, D. (2001). *Surveys in social research*. New York, NY: Routledge.

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, 2(2), Article 4. doi:10.5206/cjsotl-rcacea.2011.2.4

Dochy, F. (2009). The edumetric quality of new modes of assessment: Some issues and prospects. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 85–114). New York, NY: Springer Science+ Business Media B.V.

Donald, J. G. (1997). Higher education in Quebec: 1945–1995. In G. A. Jones (Ed.), *Higher education in Canada: Different systems, different perspectives* (pp. 161–188). New York, NY: Garland.

Dysthe, O., Engelson, K. S., Madsen, T. G., & Wittek, L. (2008). A theory-based discussion of assessment criteria: The balance between explicitness and negotiation. In A. Havnes & L. McDowell (Eds.), *Balancing dilemmas on assessment and learning in contemporary education* (pp. 121–131). New York, NY: Routledge Taylor & Francis Group.

Ecclestone, K., & Swann, J. (1999). Litigation and learning: Tensions in improving university lecturers' assessment practice. *Assessment in Education*, 6(3), 377–389.

Furedi, F. (2010). Introduction to the marketisation of *higher education and the student as consumer*. In M. Molesworth, R. Scullion, & E. Nixon (Eds.), *The marketisation of higher education and the student as consumer* (pp. 1–8). London, UK: Routledge.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London, UK: Farmer Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487

Joughin, G. (2009). Introduction: Refocusing assessment. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 1–11). New York, NY: Springer Science+Business Media B.V.

Leach, L. (2012). Optional self-assessment: Some tensions and dilemmas. *Assessment & Evaluation in Higher Education*, 37(2), 137–147. doi:10.1080/02602938.2010.515013

Luth, R. (2010). Assessment and grading at the University of Alberta: Policies, practices, and possibilities. A report to the Provost and the University. Retrieved from <https://uofa.ualberta.ca/-/media/ualberta/centre-for-teaching-and-learning/ctl-reports/assessment-and-grading-at-the-university-of-alberta-policies-practices-and-possibilities-june2010.pdf>

Morgan, C., Dunn, L., Parry, S., & O'Reilly, M. (2004). *The student assessment handbook*. London, UK: Routledge Falmer.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35(3), 277–289. doi:10.1080/02602930903541007

Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed.). London, UK: Routledge.

Ricketts, C., & Wilks, S. J. (2002). Improving student performance through computer-based assessments: Insights from recent research. *Assessment & Evaluation in Higher Education*, 27, 475–479. doi:10.1080/026029302200009348

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.

Student Participation in Quality Scotland (sparqs). (2013a). *Celebrating student engagement: Successes and opportunities in Scotland's university section*. Edinburgh, UK: Author. Retrieved from <http://www.sparqs.ac.uk/upfiles/UNI%20CELEB%20REPORT%20SPREADS%20FINAL%20.pdf>

Student Participation in Quality Scotland (sparqs). (2013b). *A student engagement framework for Scotland*. Edinburgh, UK: Author. Retrieved from <http://www.sparqs.ac.uk/upfiles/SEFScotland.pdf>

Suskie, L. (2009). *Assessing student learning*. San Francisco, CA: Jossey-Bass.

Terrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education in Practice*, 6(6), 354–363. doi:10.1016/j.nedt.2006.07.006

Thelwall, M. (2000). Computer-based assessment: A versatile educational tool. *Computers & Education*, 34, 37–49. doi:10.1016/S0360-1315(99)00037-8

Trowler, P., & Bamber, R. (2005). Compulsory higher education teacher training: Joined-up policies, institutional architectures and enhancement cultures. *International Journal for Academic Development*, 10(2), 79–93. doi:10.1080/13601440500281708

van der Velden, G. (2012). Institutional level student engagement and organizational cultures. *Higher Education Quarterly*, 66(3), 227–247.

Weigle, S. (1999). Investigating rater/prompt interaction in writing assessment: Qualitative and quantitative approaches. *Assessing Writing*, 6(2), 145–178.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Contact Information

Cheryl-Anne Poth
Centre for Research in Applied Measurement and Evaluation
Department of Educational Psychology
Faculty of Education
cpoth@ualberta.ca

Cheryl Poth is currently Associate Chair (Undergraduate Programs) and a faculty member of the Centre for Research in Applied Measurement and Evaluation, within the University of Alberta's Department of Educational Psychology. She brings experience as an international and domestic classroom teacher to her instructor and past coordinator roles (2009–2013) in the required undergraduate classroom assessment course of the pre-service teacher education program. She also serves as an assessment advisor on university, provincial and national boards.

Alex Riedel is currently a teacher in a secondary school in Germany. Prior to starting teaching, he graduated with a Master's of Education from the Measurement, Evaluation and Cognition program within the Department of Educational Psychology at the University of Alberta. During his graduate studies and work, he focused on the development of questionnaires and assessment in higher education. Before receiving his MEd, he graduated from the University of Erlangen-Nuremberg in Germany with a teaching degree.

Robert Luth has been a faculty member at the University of Alberta since 1989 and is currently a professor in the Department of Earth and Atmospheric Sciences. His primary research interests are in mantle geochemistry and petrology. He has a strong interest in education, especially at the undergraduate level, and served as associate chair (Undergraduate Programs) in his department from 1998–2004 and 2008–2013. He has been active in the university's governance since 2000 and was a Provost's Fellow in the Office of the Provost and Vice-President (Academic) in 2010–11, then Vice-Provost (Academic Programs and Instruction) in 2014–2015.